Language Variation and Change, **21** (2009), 437–453.

doi:10.1017/S0954394509990159

On the role of vowel duration in the New Zealand English front vowel shift

CHRISTIAN LANGSTROF
University of Freiburg

ABSTRACT

This article investigates the role of vowel duration in the front vowel system of New Zealand English (NZE), drawing on data obtained from speakers born between the 1890s and the 1930s. After providing a brief overview of the history of short vowels in NZE, a comprehensive analysis of front vowel duration in conjunction with a number of earlier results from formant frequency measurements will be presented. It will be shown that the front vowel system of NZE shows interaction between vowel duration and formant frequency. A number of implications that follow from these patterns for the front vowel system of NZE will be discussed. It will be argued that it is reasonable to divide up the class of short front vowels in NZE into a short set (consisting only of one vowel) and a "not-so-short set." In addition, it will be concluded that phonological class membership is irrelevant to making generalizations over patterns of movements in vowel change.

This article investigates the role of vowel duration in the front vowel system of New Zealand English (NZE). Duration has largely been ignored in previous studies on the development of the short front vowels (SFVs) in NZE. In the present article, it will be argued that vowel duration plays an integral role in the process of the restructuring of the NZE front vowel space over the last 150 years. It will be shown that it is reasonable to divide up the class of SFVs (/Iɛ æ/) in NZE into a "short set" (consisting only of one vowel) and a "not-so-short set," for want of a more appropriate designation. In the context of the NZE chain shift that involves the SFVs, duration will be shown to constitute a crucial phonetic parameter in resolving vowel distributions that overlap in F1/F2 space in intermediate stages of the shift. Therefore, the existence of solid length differences in these intermediate stages of the chain shift might have contributed to there being a chain shift in the first place rather than a merger.

I would like to thank the *Deutscher Akademischer Austauschdienst* as well as the University of Canterbury for funding the research that is discussed here. I would also like to thank Dr. Jen Hay, Professor Lyle Campbell, Dr. Margaret Maclagan, Professor Peter Trudgill, and Professor Elizabeth Gordon for their support and their insightful comments on my work. I am also indebted to Elizabeth Gordon for allowing access to the Origins of New Zealand English (ONZE) data. The intermediate archive data was collected by Rosemary Goodyear, Lesley Evans, and members of the ONZE team. The work done by members of the ONZE Project in preparing the data, making transcripts, and obtaining background information is also acknowledged.

In addition to this, the existence of solid durational differences between different short vowels might help elucidate the problem of the "erratic" behavior of these vowels in the theoretical framework on chain shifts proposed by Labov (1994). In Labov's view, the NZE SFV chain shift constitutes a counterexample to the general observation that short vowels generally fall in chain shifts, that is, show downward movements in vowel space over time. The data presented here suggests that two of the vowels in question (namely /E/ and /æ/) can be regarded as somewhat different from a "true" short vowel as /I/ in terms of their durational properties. Because these vowels are not particularly good examples of short vowels, it follows that their behavior is indeed in accordance with the Labovian regularities in chain shifts. It will be argued that what counts in predicting the pathways of vowels in chain shifts are the phonetic properties of vowels rather than their phonological status.

This article is structured as follows. The Background section provides a brief overview of the history of short vowels in NZE. The Methods section discusses the data sample. The Results section provides a comprehensive analysis of front vowel duration in conjunction with a number of earlier results from formant frequency measurements. It will be shown that, on a phonetic level at least, the front vowel system of NZE shows an interaction between vowel duration and formant frequency. The Summary/Discussion and Conclusions section discuss a number of implications that follow from these patterns for the front vowel system of NZE.

BACKGROUND

The repartitioning of the NZE front vowel space is now one of the most widely documented changes in progress in contemporary English (cf. Bauer, 1979, 1986, 1992; Gordon, Campbell, Hay, Maclagan, Sudbury, & Trudgill, 2004; Maclagan, 1982; Trudgill, Gordon, & Lewis, 1998). It mainly includes a vowel shift between the short vowels /I/, /ɛ/ and /æ/. In addition, it has been shown that correlations exist between the quality and duration of these vowels and a number of other vowels in the front vowel system of NZE, namely /i/ (Maclagan & Hay, 2007) and /a/ (Gordon et al., 2004). I will refer to the vowels in question by their lexical set names, where KIT = /I/, DRESS = /ɛ/, and TRAP /æ/ (cf. Langstrof 2006a:ch. 1 for discussion; Wells, 1982).

The SFV shift itself consists of raising and fronting of the lower two vowels TRAP and DRESS as well as the centralization and lowering of the highest vowel KIT, which results in a restructured short vowel space consisting of two front vowels (DRESS and TRAP), two central vowels (KIT and STRUT), and two back vowels (FOOT and LOT). Although it has been shown that raised variants were already present in the speech of the earliest European immigrants to New Zealand (Gordon et al., 2004), most of the chain shift is endemic (Langstrof, 2006a:ch. 3). In addition, it is now clear that the raising of TRAP and DRESS predated the centralization of KIT, which implies that the process constitutes a push-chain (Langstrof, 2006b).

https://doi.org/10.1017/S0954394509990159 Published online by Cambridge University Press

Although most studies have focused on the SFV shift in isolation (and therewith implicitly regard this as a self-contained historical process), there is evidence of interaction between the SFVs and other vowels in NZE. Gordon et al. (2004) showed that there exists, for some speakers, a significant correlation between the short vowel TRAP and the long vowel /a/ (henceforth START) in their data from first- and second-generation New Zealanders. Specifically, for the speakers for whom the correlation holds, frontness in START correlates positively with frontness and closeness in TRAP. However, it is unclear whether this constitutes a truly causal relationship between the two elements, because there are essentially three ways for the observed correlation to come about: (1) via a push-chain relationship, which would imply that fronting of START triggered the entire short vowel shift; (2) as a secondary pull-chain effect whereby START moves into the vowel space vacated by raised TRAP; (3) no causal relation at all, which means that the correlation comes about incidentally as speakers of comparable innovativeness in a given vowel change often show a similar degree of innovativeness in another one.

A more in-depth analysis of an interrelation between a short vowel and a long vowel is provided by Maclagan & Hay (2007), who analyzed high front vowels in contemporary NZE and found both a correlation between vowels of different subsystems as well as providing evidence for trade-off relationships across different phonetic dimensions. It is shown that formerly short DRESS is now a high monophthong, which is distinguished from FLEECE mainly on the basis of diphthongization and duration rather than quality in the speech of younger informants. Although it is not unusual for mid-monophthongs to push out high monophthongs into the diphthong system (a similar process happened during the Great Vowel Shift) as a result of a raising process, the interesting facet to note in this case is the fact that DRESS is originally a "short" vowel. This process therefore reflects one of the allegedly rare cases of a clear functional relationship over time between vowels from different subsystems (cf. the extensive discussion in Labov [1994] on this topic). The obvious alternative to the preceding conclusion is to regard contemporary NZE DRESS as a long vowel, which would make the preceding process rather less mysterious. The Results section of this article presents data from earlier speakers of NZE, which suggests that this alternative view might be preferable on phonetic grounds, while the Summary/Discussion section discusses a number of problems that arise from this solution.

METHODS

The speaker sample

The speakers analyzed in the present article are New Zealanders born between the late 1890s and the mid-1930s. Thus, they represent an intermediate stage between the speakers analyzed by Gordon et al. (2004) and contemporary New Zealanders. These speakers will be referred to as "intermediate speakers." Nine speakers who represent a subset of a larger corpus of approximately 110 speakers held at the Origins of New Zealand English (ONZE) archive at the University of Canterbury were analyzed. The speakers are divided into two groups. The first group consists of four male speakers born at the beginning of the intermediate period (more specifically, between the 1890s and the early 1900s). The second group consists of four female speakers born toward the end of that period. The design of the sample was motivated by the following considerations. A parallel analysis of vowel quality of 30 intermediate speakers discussed in Langstrof (2006a, 2006b) revealed that it was in the intermediate period that the shift in the NZE SFV system gained momentum. In addition, the early male group and the late female group represent the polar ends of the innovativeness spectrum with regard to the quality change in the SFVs. It was therefore assumed that if any significant changes occurred over the intermediate period in terms of vowel duration as well, these should show up best if speakers at the opposite end of the innovativeness scale are compared. It should also be noted that the labels "early" and "late" as used in the following analysis refer to the intermediate period only rather than the overall history of the development of NZE.

The speech material consists of interviews of varying lengths that were recorded by various interviewers in the early- to mid-1990s in the course of a project on oral history in New Zealand. Subjects were asked to supply information about their childhood in various regions of New Zealand. The material on which the present study is based consists of running speech.

Analysis

Acoustic analysis of the vowels KIT, DRESS, and TRAP was carried out using Praat software as well as the plug-in Akustyk.[1] The duration measurements reported were extracted by Akustyk, whereas formant frequency measurements were taken directly in Praat. The formant frequency data were subsequently normalized using Lobanov's z-score algorithm (Lobanov, 1971) and coded for a number of independent variables (speaker age, voicing status of the following consonant, number of syllables, F1, F2) and analyzed using classification and regression tree (CART) analysis.[2] Duration measurements were taken from the visible onset of modal voicing on the spectrogram. Only stressed tokens were measured. The token numbers obtained from each speaker/vowel are given in Table 1. Table 2 shows token numbers for monosyllables and polysyllables, respectively. Only tokens before alveolar stops were measured for DRESS and TRAP. For KIT, prealveolar as well as prevelar tokens were measured.[3]

RESULTS

The duration data obtained from the nine intermediate speakers are shown in Table 3.

TABLE 1. *Number of tokens obtained from two groups of intermediate speakers*

		KIT		DRESS		TRAP	
Group	Speaker	_[+voice]	_[–voice]	_[+voice]	_[–voice]	_[+voice]	_[–voice]
Early	1	17	22	6	2	7	5
	2	7	17	4	4	5	6
	3	9	17	7	6	10	6
	4	11	6	5	3	6	5
	Total	44	62	22	15	28	22
Late	1	7	9	5	4	6	5
	2	5	6	3	2	3	5
	3	8	5	5	7	5	6
	4	5	7	1	1	1	5
	5	4	3	0	1	3	5
	Total	29	30	14	15	18	26

[+/–voice] specifies the voicing state of the following consonant.

Vowels in monosyllables

The data given in Table 3 suggest a number of pertinent points. First of all, as can be seen in the second column, both DRESS and TRAP are appreciably longer than KIT in the speech of the intermediate speakers, where the overall ratios are approximately 1:2 for KIT VS. DRESS, approximately 1:2.1 for KIT VS. TRAP, and 1:1.05 for DRESS VS. TRAP in monosyllables. The duration difference between KIT and DRESS is significant (Wilcox test, $W = 442.5$, $p < .0001$), but the difference between DRESS and TRAP is not. The same pattern holds if we break down the data into the two age groups. The differences between the age groups are not significant. In addition, the well-known length difference within a vowel depending on whether it occurs before a voiced stop or a voiceless stop holds within the present sample (where the difference is significant at $p < .0001$ [$W = 1607.5$] for the KIT vowel, at $p < .0001$ for DRESS [$W = 385.5$], and at $p < .01$ for TRAP [$W = 776.5$]).

TABLE 2. *Token numbers obtained from two groups of intermediate speakers broken down into occurrences in monosyllabic and polysyllabic words*

		Syllabicity		
Group	Vowel	Mono	Poly	Ratio
Early	KIT	61	45	1.35
	DRESS	24	13	1.84
	TRAP	32	18	1.77
Late	KIT	32	27	1.18
	DRESS	21	8	2.62
	TRAP	34	10	3.40

https://doi.org/10.1017/S0954394509990159 Published online by Cambridge University Press

TABLE 3. *Average durations in milliseconds (ms) of SFVs in the speech of 9 intermediate speakers, 4 early males, and 5 late females*

	Monosyllables			Polysyllables		
	KIT	DRESS	TRAP	KIT	DRESS	TRAP
All	49.65	99.37	106.94	46.64	75.57	93.46
Early	51.66	104.1	103	49.27	74.08	94.44
Late	45.81	94.38	110.65	42.43	78	91.70
_[+voice]	59.23	115.3	118.7	46.32	76	97.75
_[–voice]	42.72	72.44	97.17	46.92	73.29	87.75
Early _[+voice]	62.58	114.4	112.1	48.39	77.33 ($n = 3$)	96.73
Late _[+voice]	52.54	116.9	127.2	44	82.25 ($n = 4$)	100
Early _[–voice]	43.54	69.2	92.67	49.85	73.1	90.86
Late _[–voice]	42.21	73.91	100.4	40.33	73.75	83.40 ($n = 4$)

Token numbers below $n = 5$ are indicated.

If the data are broken down further into age + voicing of the following consonant, the duration difference between KIT vowels before voiced and voiceless consonants fails to reach significance in the late sample. The differences in DRESS and TRAP retain significance within both groups.

Overall, this accords with the internal partitioning within the SFV set with respect to the chain-shift in F1/F2 space, in that it is DRESS and TRAP that underwent an upward movement, whereas KIT centralized. Here we have the same subgrouping (KIT VS. DRESS and TRAP).

Vowels in polysyllables

If we compare the durations of the SFVs in monosyllables to those in polysyllables, it seems clear that whereas DRESS and TRAP are shorter in polysyllables, this does not hold for KIT. The duration difference between monosyllables and polysyllables is significant for DRESS only ($W = 276.5$, $p < .01$). In addition, it is only in polysyllables where the duration difference between DRESS and TRAP is significant ($W = 153.5$, $p < .01$). The contrast in duration between prevoiced and prevoiceless tokens fails to reach significance in all vowels, which might be partly due to low token numbers.

It seems that the results quoted as to the effect of polysyllabicity on duration are reflected to some extent in the present analysis (i.e., the DRESS vowel). In addition, there is some support for Klatt's (1973) assumption that there is no linear "adding-up" of shortening factors but rather some minimal level of vowel duration around which shortening factors do not apply anymore. This is exemplified by the KIT vowel, where duration differences are minor and reach significance only between two conditions: (1) early male monosyllables closed by a voiced stop vs. early male monosyllables closed by a voiceless stop ($W = 733$, $p < .0001$), and

(2) early male monosyllables closed by a voiced stop vs. early male polysyllables closed by a voiced stop ($W = 348.5$, $p < .01$).

On the whole, it seems that shortening factors do indeed interact. Figure 1 shows duration averages of all three vowels before voiced and voiceless stops in monosyllables (Figures 1a and 1b) and polysyllables (Figures 1c and 1d). It seems that for a vowel to be subject to shortening factors, it needs to have a certain minimum length. That is, the only significant shortenings occur in DRESS/TRAP where monosyllabic tokens before voiced stops undergo similar degrees of shortening in both polysyllables as well as in voiceless environments. In addition, these shortening factors do not stack. The KIT vowel is, by and large, exempt from either process with the exception of early male KIT before voiced stops, which undergoes shortening. As far as duration averages are concerned, we conclude that DRESS and TRAP are not only significantly longer than KIT, but also have a number of properties typical of long vowels such as susceptibility to shortening in certain environments as well as a greater internal (i.e., within one vowel) durational variability.

Factor analysis and the relation between duration and formant frequency

A CART analysis was carried out in order to clarify the relation between duration and formant frequency as it holds in the intermediate speaker sample.[4] This was

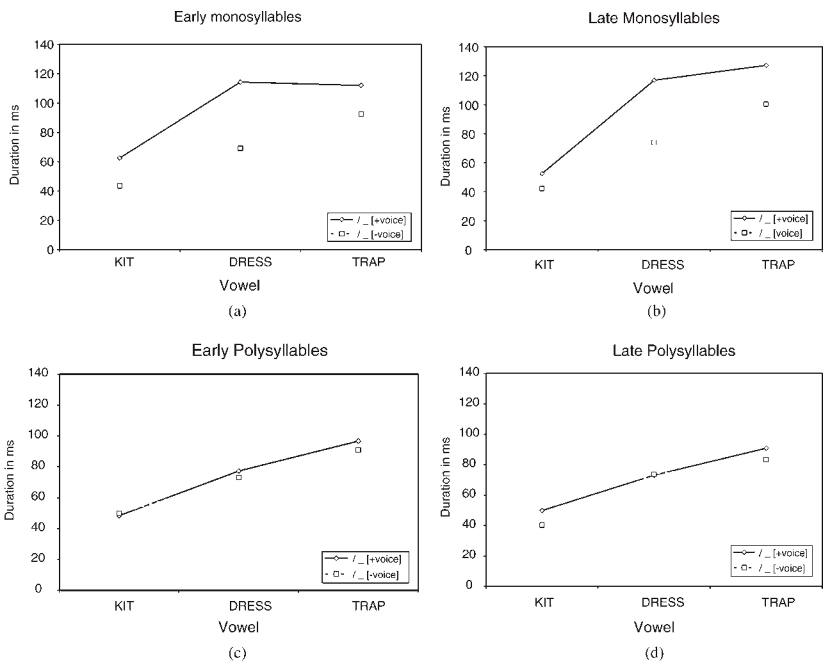

FIGURE 1. Average durations of KIT, DRESS, and TRAP before voiced and voiceless stops: monosyllables in the early male (a) and early female (b) samples; polysyllables in the early male (c) and early female (d) samples.

https://doi.org/10.1017/S0954394509990159 Published online by Cambridge University Press

done to test for the solidity of the subgroupings inferred from the data discussed previously. If DRESS and TRAP are truly different than KIT in terms of duration, this should be reflected in a regression analysis. In addition, this subgrouping should be relatively more prominent compared with other factors that are known to have an impact on vowel duration, such as voicing of the following consonant or syllabicity.

Figure 2 shows the result of a CART analysis carried out by taking into account a further continuous factor, namely formant frequency. In Figure 2, duration is the dependent variable regressed onto the following independent variables: vowel (*KIT/DRESS/TRAP*), F1, F2, age group (*early* vs. *late*), syllable (*mono*[syllabic]/*poly* [syllabic]), voicing status of the following consonant (*voiced/voiceless*).

As was to be expected on the grounds of the data presented, the main predictor of overall duration is vowel identity, whereby KIT is appreciably shorter than DRESS and TRAP. In addition, no further factors partition the KIT branch, whereas the DRESS/TRAP branch is split into vowels before voiceless stops (average duration 85.29 ms) and voiced stops (average duration 110.7 ms). This distinction takes precedence over vowel identity, which is a factor only within the *voiceless* node, where DRESS is shorter than TRAP. For DRESS/TRAP vowels before voiced stops, it is syllable structure that predicts duration in that both DRESS and TRAP tend to be markedly longer if they occur in monosyllables closed by a voiced stop than in polysyllables closed by a voiced stop. The two major conclusions from the Analysis section therefore seem reflected in the pooled CART data as well. "Shortness" of a vowel goes along with a lack of further internal partitioning along the duration dimension (as KIT represents a terminal node, whereas DRESS/TRAP do not) and shortening factors interact, although with respect to Figure 2, we may prefer to say "lengthening factors interact" because the terminal branching that relates to a phonetic lengthening factor (i.e., syllabicity) stems

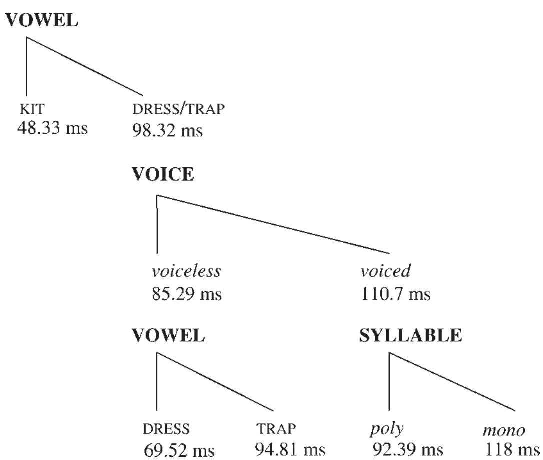

FIGURE 2. Classification and regression tree of pooled duration measurements along the independent variables F1, F2, age, vowel, voice, and syllable.

https://doi.org/10.1017/S0954394509990159 Published online by Cambridge University Press

from the superordinate node, which has the higher duration average (i.e., the DRESS/TRAP tokens before voiced stops). Before voiceless stops, vowel identity is a better predictor of duration (where DRESS is shorter than TRAP).

What is striking is that in the pooled data, formant frequency does not show up as a factor in predicting vowel length, although this would be expected on the grounds that F1 has frequently been found to correlate positively with duration (cf. Peterson & Lehiste, 1960). In addition, "peripherality" has previously been related to length (cf. Labov, 1994; Labov, Yaeger, & Steiner, 1972, for a discussion of the relevant terminology) to the effect that the more peripheral a vowel, the longer it would be expected to be. Because we are concerned with front vowels, we would therefore expect a positive correlation between F2 and duration.

This interaction does indeed show up if we break up the overall data pool by vowel. Figures 3a to 3c show separate CART trees for the three vowels. The variables are as previously given. The trees are truncated below the second branching.

Within both the KIT and the DRESS trees, the major predictive factor in determining duration is the voicing status of the following consonant. Vowels before voiced consonants are longer than before voiceless ones (where the ratios are 1:1.2 for KIT and 1:1.58 for DRESS). Two points deserve mention in this context. First of all, it is somewhat surprising to find this division in the KIT data as it seems to be mainly a property of the early group, which implies that if anything, we would expect this as a subordinate branching under a superordinate branching along the factor age. On the other hand, the late group does show a length difference (albeit not significant) within KIT to the same effect, which suggests that this difference is carried by the late group in the pooled KIT data. Second, it looks as if the duration difference of vowels before voiced vs. voiceless consonants which we noted as a property of the overall DRESS/TRAP data in Figure 3 was carried mainly by DRESS, as it fails to show up as a predictive factor in the TRAP data (Figure 3c). In the KIT data, second formant frequency is a subordinate factor below both primary branches. The data are somewhat contradictory because within the group of tokens before voiceless consonants, high F2 values seem to predict lower vowel duration, whereas within the sample of vowels that occurs before voiced stops, the inverse is the case. Although this seems unexpected, the vowel tokens for which the negative correlation holds constitute the shortest subsample overall. If we recall the notion as outlined previously that the longer the intrinsic duration of a vowel, the more well-established the phonetic mechanisms which favor lengthenings/shortenings, this is probably the subsample where overriding of the expected mechanisms is least unexpected.

The TRAP data are best predicted by second formant frequency in that the higher the F2, the longer the duration of a vowel token. In addition, we note in both the DRESS and the TRAP data an effect of syllable structure whereby vowels in monosyllables have higher duration averages than vowels in polysyllables within a branching that is already of comparatively high average duration.

https://doi.org/10.1017/S0954394509990159 Published online by Cambridge University Press

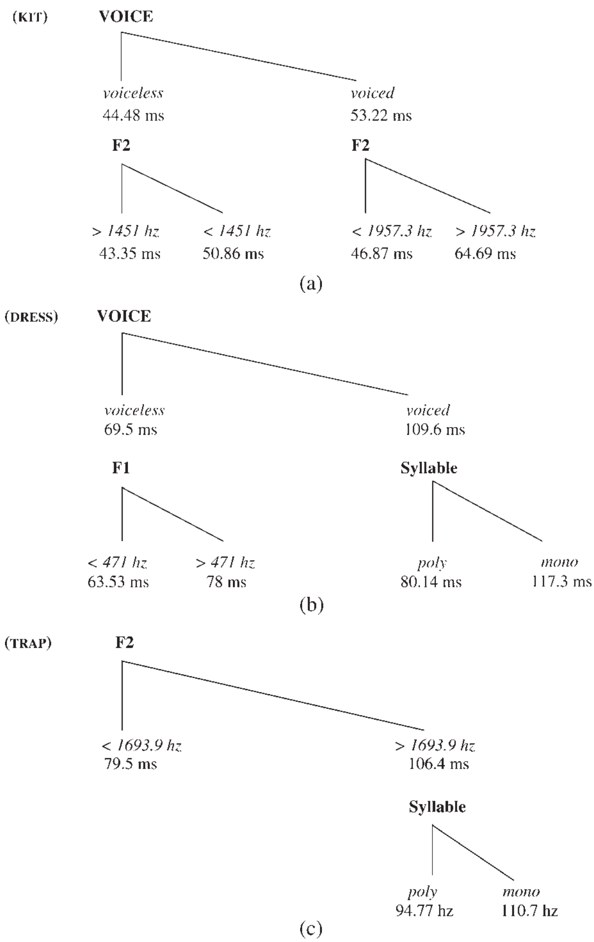

FIGURE 3. Classification and regression tree analysis of duration measurements broken down into the three SFVs.

It is interesting to note that the expected correlation between F1 and duration shows up only within one terminal branch and for one vowel (DRESS) only. In addition, it fails to reach significance in a Spearman correlation test ($\rho = .27$, $S = 2955$, $p > 1$). Overall correlation test results are summed up in Table 4.

It seems clear that there is nowhere a statistically significant correlation between F1 and vowel duration. However, some F2 effects can be noted. In the KIT data, there is a positive correlation between F2 and duration in monosyllables and before voiced stops, which is in keeping with the preceding analysis. Within the DRESS/TRAP data, F2 correlates with duration overall, although the effect seems to be carried by monosyllables only. In addition, there is an interesting mismatch between the two vowels in terms of how they behave across the two speaker

https://doi.org/10.1017/S0954394509990159 Published online by Cambridge University Press

TABLE 4. *Correlation coefficients between the variables* duration *and* F1/F2 *in the SFVs as obtained by a Spearman rank correlation (where ρ = correlation coefficient)*

		Duration				Duration	
F1		ρ	p Value	F2		ρ	p Value
KIT	all	–0.11	>.1	KIT	all	0.12	>.1
	early	–0.11	>.3		early	0.1	>.2
	late	0.03	>.7		late	0	>.9
	mono	–0.14	>.1		mono	**0.2**	**<.5**
	poly	–0.06	>.6		poly	0	>.9
	/_[+voice]	–0.18	>.1		/_[+voice]	**0.39**	**<.001**
	/_[–voice]	0.1	>.3		/_[–voice]	–0.13	>.2
DRESS	all	–0.16	>.1	DRESS	all	**0.34**	**<.01**
	early	–0.22	>.1		early	**0.46**	**<.01**
	late	–0.18	>.3		late	0.21	>.1
	mono	–0.23	>.1		mono	**0.38**	**<.05**
	poly	0.25	>.2		poly	–0.03	>.8
	/_[+voice]	–0.09	>.6		/_[+voice]	0.26	>.1
	/ _[–voice]	0.27	>.1		/_[–voice]	–0.22	>.2
TRAP	all	–0.03	>.7	TRAP	all	**0.24**	**<.05**
	early	0.06	>.6		early	–0.01	>.9
	late	–0.04	>.7		late	**0.54**	**<.001**
	mono	–0.02	>.8		mono	**0.39**	**<.01**
	poly	0.03	>.8		poly	–0.19	>.3
	/_[+voice]	0.03	>.8		/_[+voice]	0.13	>.3
	/_[–voice]	0.06	>.6		/_[–voice]	**0.36**	**<.05**

Correlations that reach statistical significance below the 5% level are given in bold.

groups. Whereas duration and F2 correlate in DRESS in the early group only, the inverse is the case in the TRAP data.

Duration and F1/F2 overlap

Given the preceding results, we can conclude that DRESS and TRAP are different from a "true" short vowel such as KIT. The reasons for this development in NZE can probably be sought in the mechanisms of the SFV shift itself. There is an appreciable degree of F1/F2 overlap between KIT and DRESS in the speech of early intermediate speakers (cf. Langstrof, 2006a, 2006b). This overlap can be shown to exist in the speech of one and the same individual, which implies that additional cues are necessary to disambiguate vowel tokens in regions of overlap to prevent merger.

Figure 4a shows that for one and the same individual, a number of tokens are ambiguous between DRESS and KIT in F1/F2 space. However, if we plot length against the first formant frequency for the same speaker (Figure 4b), this overlap largely disappears. It therefore appears that there is, to some extent, a trade-off relation between duration and formant frequency in terms of overlap between adjacent vowel distributions. Figure 5 plots duration against both formant frequencies for the early male and the late female sample.

https://doi.org/10.1017/S0954394509990159 Published online by Cambridge University Press

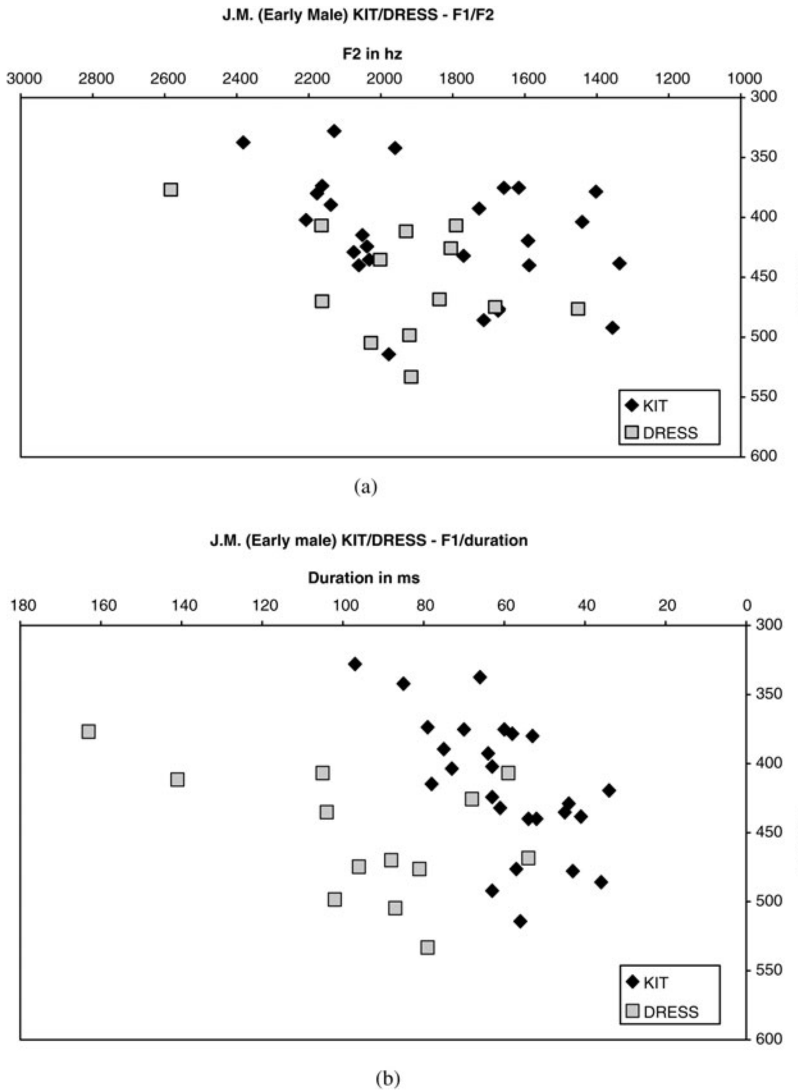

FIGURE 4. Differential degrees of overlap between the lexical sets of KIT and DRESS in the speech of an early male speaker. Figure 4a plots the first two formants against each other, and 4b plots duration against the first formant.

In the F1/F2 dimension, the early sample shows nearly complete overlap between DRESS and its neighboring vowels. This overlap is resolved over the intermediate period by KIT centralisation and DRESS raising (Figure 5b). However, if we plot frequency against duration (Figures 5c to 5f) it appears that the TRAP distribution moves toward the KIT distribution in the F1/duration plot. Table 5

https://doi.org/10.1017/S0954394509990159 Published online by Cambridge University Press

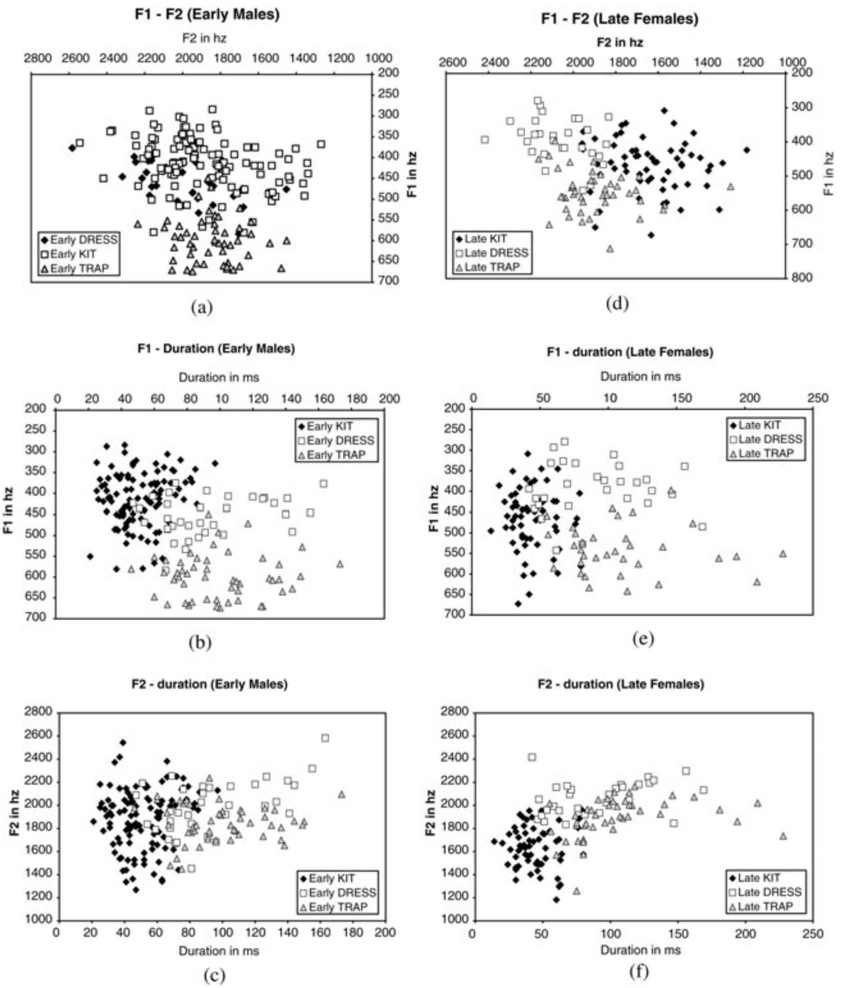

FIGURE 5. F1/F2 duration plots of KIT, DRESS, and TRAP in the early male and late female sample.

presents Hotelling-Lawley trace scores (a measure of the relative distance between distributions) for each of the distributions plotted in Figure 5.

The Hotelling-Lawley traces shown in Table 5 indicate wider separation between KIT and DRESS in the late sample compared with the early sample. As for DRESS vs. TRAP, it is only in the F1/duration dimension where the distance between the two distributions decreases over time. Finally, KIT and TRAP approximate each other in all dimensions.

The assumption that speakers/listeners of a language that is undergoing a vowel change whereby the distributions of two vowel phonemes approximate each other over time take advantage of (or "exaggerate") alternative cues to the identity of a token in the overlap region seems plausible. What seems to be important here is that

TABLE 5. *Hotelling-Lawley trace scores of the distributions of* KIT, DRESS, *and* TRAP *in the early male and the late female sample*

		F1–F2	F1 Duration	F2 Duration
Early Males	KIT—DRESS	0.17	0.97	0.79
	DRESS—TRAP	1.99	2.07	0.2
	KIT—TRAP	2.11	3.58	1.33
Late Females	KIT—DRESS	1.53	1.07	1.99
	DRESS—TRAP	1.6	1.6	0.4
	KIT—TRAP	0.95	1.57	1.52

All scores are significant below .01.

subphonemic durational cues can be dialect-specific and induced by historical processes such as vowel shift. In addition, the analysis presented earlier suggests that they are part of an extrinsic allophony and are therefore variable in time and space.

What requires an explanation is why there should have occurred a reactionary movement of KIT, given that in this intermediate situation the two vowels were still solidly distinguished in F1 vs. duration space. The only hypothesis I have to offer here is purely conjectural. During the stage of the shift as exemplified by speakers such as J.M., we find three vowels in rather close proximity to each other, namely KIT, DRESS, and FLEECE. What this amounts to is a three-way distinction in vowel length in that region of the vowel space, which is not mirrored anywhere else in the vowel system. It has been pointed out before (cf. Bybee, 2001, on the rise and fall of early Middle English rounded front vowels) that although such constellations can arise as the outcome of nonteleological sound change, they seem to be abandoned rather quickly. That is, phoneme systems tend to have similar types of distinctions across the board. In the case of the SFV shift in the intermediate period, this means that the duration difference might be best regarded as a "transitional cue" in disambiguating F1/F2 overlap between two phonemes.

SUMMARY/DISCUSSION

The important points to note are as follows:

- DRESS and TRAP are considerably longer than KIT.
- For all three lexical sets, vowels before voiced consonants are longer than before voiceless ones.
- In addition, this effect is stronger in DRESS than it is in any of the other two vowels.
- Lengthening factors combine more linearly in longer segments.
- There is no statistically significant effect of F1 (i.e., height) on vowel length.
- In monosyllables, there is a strong effect of F2 on vowel length, whereby on the whole, frontness correlates positively with length.
- TRAP differs from the other two vowels in that the F2 effect is more important in predicting vowel duration than the voicing state of the following consonant.

- There are no statistically significant duration differences between the two speaker groups in any of the SFVs.
- Duration is a solid predictor of vowel identity in overlapping distributions.

From these results, a number of important implications follow with regard to front vowel length in the intermediate period. First of all, overall vowel length indicates that DRESS and TRAP are in fact different from KIT, because they are appreciably longer. Against the background of the front vowel shift, we can therefore conclude that they are indeed “different things” on the grounds of their differential behavior along two phonetic dimensions (i.e., DRESS and TRAP both raise/front *and* are longer; KIT centralizes and is shorter).

Second, the lack of any strong correlation between F1 and length permits two possible interpretations of the relation between height and quantity in NZE. The straightforward one would be to conclude that in NZE, intrinsic constraints on vowel length depending on the vowel’s openness are simply overridden. The appeal of this assumption probably depends on how mechanistically one interprets such a constraint in the first place, that is, whether this relation is primarily language- or dialect-specific, or whether it follows from motoric constraints in speech production. Under the latter premise, this interpretation of vowel length in NZE becomes rather ad hoc and mysterious. The alternative would be to assume that the condition holds, but that there has been moderate subphonemic lengthening over time in the SFV set in NZE that offsets this intrinsic effect and thereby preserves isometry in the face of movement across articulatory space. However, the data presented here permit no clear-cut conclusion in this regard.

With respect to Labov’s (1994) theory of directional vowel shift, the present analysis casts doubt upon the alleged exceptional status of DRESS and TRAP in NZE. Phonetic shortness of DRESS and TRAP is hard to ascertain in the face of the presence of a phonetically much shorter vowel, namely KIT. Apart from phonetic duration, it was also shown that DRESS and TRAP show a number of phonetic interactions that are more typical of long vowels, such as a larger duration difference depending on whether the vowel occurs in mono- or polysyllables or before a voiced or voiceless stop. It may therefore be advisable to remain cautious in positing that DRESS and TRAP are short vowels in NZE and should therefore be expected to behave like LOT, STRUT, or FOOT rather than their tense counterparts. Under this interpretation, the behavior of DRESS and TRAP is therefore well in line with the general principles of vowel shifts.

CONCLUSIONS

In this article, I showed that there exist solid duration differences between lax front vowels in NZE and argued that this duration difference might have been the crucial factor that allowed listeners to keep the distributions of otherwise overlapping vowels apart. Labov & Baranowski (2006) came to a similar conclusion. In a perceptual study of vowel contrasts in American English, it was found that

https://doi.org/10.1017/S0954394509990159 Published online by Cambridge University Press

a durational difference of 50 ms is sufficient for listeners to reliably keep vowel distributions apart in perception. It therefore seems advisable to take a closer look at the role of vowel duration in studies on vowel change in the future.

In the case of the speakers analyzed in the present article, both DRESS and TRAP are considerably longer than KIT. This subgrouping reflects the patterning of these vowels in the NZE front vowel shift, where DRESS and TRAP undergo raising and fronting, whereas KIT undergoes centralization. In addition, only DRESS and TRAP show a considerable degree of durational allophonization, a property usually found in long vowels. We can therefore justifiably posit that these elements may be different from "true" short vowels on phonetic and historical grounds. Taking this argument one step further, it seems advisable to separate the phonological status of vowels from their classification as phonetic and diachronic entities. Although this is probably a rather commonplace assumption, the conclusion with respect to what counts in sound change may be less so. If the subgrouping of these elements with regard to their movements in vowel spaces reflects their phonetic patterning rather than phonological class-membership (after all, all three vowels show the well-known phonotactic properties of lax vowels in English), we cannot make statements such as "tense vowels rise" and the like in the absence of a workable a priori definition of tenseness in phonetic terms. Rather, generalizations over the movements of certain elements over time in vowel space need to be based on their phonetic properties, which furthermore need to be reassessed in each variety or language. Assigning a given element (say the DRESS vowel) to a "short" vowel set on the basis of etymology a priori does not yield the correct results with regard to how one would expect a given element to behave in vowel change.

This view is corroborated by the data in the Results section, where it was shown that duration can serve as a cue toward separating vowel distributions that overlap in F1/F2 space, even in cases where the two distributions in question belong to the same subset of vowels. Therefore, what counts as a predictor in making generalizations over vowel movements are the phonetic properties of the given vowels at specific points in time rather than their phonological properties or their etymological origins.

NOTES

1. Praat: Doing Phonetics by Computer was developed by Paul Boersma and David Weenink. It is available at: www.praat.org. The Akustyk add-on was developed by Bartlomiej Plichta. It is available at: bartus.org.

2. An in-depth mathematical description of CART can be found in Breiman, Friedman, Olshen, & Stone (1984). Cf. also Mendoza-Denton, Hay, & Jannedy (2003) for a discussion regarding the advantages of CART for analyses that involve continuous dependent variables such as formant frequency values.

3. This was done because prevelar KIT tokens were shown to be the least innovative allophones in the F1/F2 dimension during the SFV shift (Langstrof, 2006a). However, no significant durational difference between prevelar tokens and prealveolar tokens showed up in the subsample discussed here.

4. Note that in the CART figures as shown throughout this article, the first branching corresponds to the statistically most significant division within the data. The CART trees always divide the data pool into two parts that reflect the most significant partitioning of the data within each node. For example, the highest-order node in Figure 3 is to be read as follows. Within the overall data pool on vowel duration, the most significant predictor of duration is vowel identity. The group of KIT tokens has an average duration of 48.33 ms, the group of DRESS and TRAP tokens has an average duration of 98.32 ms. The length of the branches is not related to the significance of the division.

REFERENCES

Bauer, Laurie. (1979). The Second Great Vowel Shift. *Journal of the International Phonetic Association* 9.2:57–66.

_______ (1986). Notes on New Zealand English phonetics and phonology. *English World-Wide* 7:225–258.

_______ (1992). The Second Great Vowel Shift revisited. *English World-Wide* 13.2:253–268.

Breiman, Leo, Friedman, Jerome, Olshen, Richard A., & Stone, Charles. (1984). *Classification and regression trees*. New York: Chapman and Hall.

Bybee, Joan. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Gordon, Elizabeth, Campbell, Lyle, Hay, Jennifer, Maclagan, Margaret, Sudbury, Andrea, & Trudgill, Peter. (2004). *New Zealand English: Its origin and evolution*. Cambridge: Cambridge University Press.

Klatt, Dennis. (1973). Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America* 55:1102–1104.

Labov, William. (1994). *Principles of linguistic change,* Vol. 1*: Internal factors*. Malden, MA: Blackwell.

Labov, William, Yaeger, Malcah, & Steiner, Richard. (1972). *A quantitative study of sound change in progress*. Report on National Science Foundation Contract NSF-GS-3287, University of Pennsylvania.

Labov, William, & Baranowski, Maciej. (2006). 50 milliseconds. *Language Variation and Change* 18:1–18.

Langstrof, Christian. (2006a). *Vowel change in New Zealand English—Patterns and implications*. Ph.D. dissertation, University of Canterbury.

_______ (2006b). Acoustic evidence for a push-chain shift in the intermediate period of New Zealand English. *Language Variation and Change* 18:141–164.

Lobanov, Boris. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America* 49:606–608.

Maclagan, Margaret. (1982). An acoustic study of New Zealand English vowels. *The New Zealand Speech Therapists' Journal* 37(1):20–26.

Maclagan, Margaret, & Hay, Jennifer. (2007). Getting fed up with our feet: Contrast maintenance and the New Zealand English "short" front vowel shift. *Language Variation and Change* 19:1–25.

Mendoza-Denton, Norma, Hay, Jennifer, & Jannedy, Stefanie. (2003). Probabilistic sociolinguistics: Beyond variable rules. In R. Bod, J. Hay, & S. Jannedy (eds.), *Probabilistic Linguistics*. Cambridge, MA: MIT Press.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32:693–703.

Trudgill, Peter, Gordon, Elizabeth, & Lewis, Gillian. (1998). New dialect formation and Southern Hemisphere English: The New Zealand short front vowels. *Journal of Sociolinguistics* 2(1):35–51.

Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.